DOCUMENT RESUME

ED 279 670 TM 870 057

AUTHOR Cohen, Michael
TITLE Designing State Assessment Systems.
PUB DATE Aug 86
NOTE 26p.; One of 46 papers commissioned by the Study Group on the National Assessment of Student Achievement and cited in Appendix B to their final report "The Nation's Report Card" (TM 870 049). For other papers in this group, see TM 870 050-094.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Quality; Educational Testing; Elementary Secondary Education; Evaluation Utilization; Federal State Relationship; Government Role; *Information Needs; *Measurement Objectives; National Surveys; Program Design; *State Programs; State Surveys; *Test Construction; *Testing Programs
IDENTIFIERS *National Assessment of Educational Progress

ABSTRACT

Guidelines for developing an ideal statewide educational assessment system are presented as one response to the redesign of the National Assessment of Educational Progress (NAEP). Emphasis is on the contribution NAEP can make to meet state needs. Core information which is useful to most states is discussed, assuming it will be directed toward policymakers and local educators. The state testing program should support the development and revision of educational policy, mobilize and sustain political support, and improve educational quality. This improvement may be accomplished by increasing the awareness of objectives, setting performance standards, providing incentives, and focusing accreditation procedures on quality. The testing program should focus on the school site level, include indicators in addition to academic achievement, rely on multiple measures, evolve over time, and focus on changes over time. The key components of useful assessment systems are annual school data profiles and a set of quality indicators. Suggested indicators which reflect long-term goals and effective schooling are listed. Public opinion, data collection, and state-by-state comparisons are also briefly discussed. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED279670

Designing State Assessment Systems

Michael Cohen

National Association of State Boards of Education

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. Cohen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper commissioned by

THE STUDY GROUP ON THE NATIONAL ASSESSMENT OF STUDENT ACHIEVEMENT

1986

TM 870 057

BEST COPY AVAILABLE

INTRODUCTION

I've been asked to prepare this "design paper" on state assessment systems for the blue-ribbon panel charged with reviewing the National Assessment of Education Progress and proposing needed revisions for future NAEP activities. My charge is to think through what an "ideal" state system might be, unconstrained by the nature or limitations of existing state assessment systems. Presumably those responsible for the direction of the national assessment effort can use this state perspective as one basis for determining the unique contribution such a national effort can make and design it accordingly. My hope is that this vision will also stimulate discussion at the state level, as state leaders rethink their own data needs.

Context

The task of redefining the National Assessment Program is part of a broader federal effort to redesign its data collection and statistical systems. This effort comes in the midst of a major education reform movement that started in the early 1980s, and which has focused on the quality and performance of the education system and its personnel. This reform agenda appears to be continuing. Within the past several months alone, reports from the Carnegie Task Force on Education and the Economy, the Holmes Group, and the National Governors' Association have each called for a second wave of reforms, this time focused squarely on strengthening, supporting and improving the working conditions of professional educators at the school site.

Most important for our purposes here is the fact that much of the ongoing reform efforts have been state initiated, with substantial political and policy leadership provided by governors and legislators, as well as state boards of education and chief state school officers. To some extent, the attention of state political leaders may simply reflect the episodic popularity of education as an issue. However, in all likelihood it more nearly reflects several deeper long term trends, trends that suggest a strong state role is likely to persist for some time into the future.

The state share of education expenditures has been steadily increasing for some time, and by now, on average nationally, states pay for approximately 50% of education costs. In nearly every state, education is the single largest item in the state budget, and will remain so. In addition, in states as varied economically as Arkansas, Mississippi, Massachusetts, and Texas, education quality is seen as a key to broader and more enduring goals of economic development. Finally, state policy leadership in education can be understood as a consequence of some two decades of modernization of the legislative and executive branches of state government, which has enhanced state policymaking capacity generally (Doyle and Hartle, 1985). Further, the National Association of State Boards of Education is currently developing a set of recommendations to modernize state boards of education. If implemented, these recommendations will further enhance state capacity for education policymaking and, at the same time, increase the demand for assessment and other policy-relevant data.

Consequently, it is fitting that the emphasis in this paper is on state, rather than federal or national, data needs. For the analysis which follows is predicated on the assumption that a state assessment system can, and needs to, be an important part of the apparatus for state education policymaking and leadership.

Assumptions Guiding This Analysis

In light of the foregoing, the primary purpose of this paper is to sketch out the design of an assessment system which will first and foremost meet state-level needs. From this point of view, questions about the design of a national assessment effort are secondary, and are considered largely in light of the contribution a national system can make to state needs. To the extent that state systems need to be shaped to accommodate a larger national effort, the bias in this paper is that such accommodations be made when and where they will best serve state purposes.

Second, state-level information needs, in their particulars, will be state-specific and therefore quite varied among the states (Fuhrman and McDonnell, 1985). Therefore my attempt here is to think through the structure or framework of state assessment systems which can be applicable across states,

rather than to fill in the substantive details which are likely to vary somewhat from state to state.

From this perspective, the identification of a "common core" of data needs is seen as the by-product of 50 individual state decisions, rather than the result of a framework imposed nationally. This serves to underscore the substantive differences between a state and national perspective, and treats the need for a system which enables state-by-state comparisons as problematic, rather than automatic. On the other hand, where these are common substantive information needs among the states, it does seem reasonable to work towards common operational definitions, thus maximizing opportunities for both national level aggregation and multi-state comparisons.

Third, I've made the assumption that the primary investors in and users of state assessment systems will be policymakers, including governors and their staff, legislators and their staff, members of state boards of education, and chief state school officers and their staff. Though each of these play different roles in the policymaking process, recent research suggests that, in general, their information needs are more nearly similar than different (Cohen, 1985). Consequently, when thinking about the components of a state assessment system, I assume that the same information will serve all policymakers equally as well.

Fourth, I see policymakers and educators at the local level -- school board members, superintendents, and principals and teachers in individual buildings -- as extremely important and only somewhat secondary users of state assessment data, at least in an ideal system. State assessment systems should be designed in such a way as to inform and guide their local improvement efforts.

Finally, I assume that resources for a state assessment system are both finite and in competition with programmatic and service needs. Thus, within a comprehensive framework clearly defining state data needs, the task is to be as parsimonious as possible, limiting costs and reducing local respondent burden. This is despite the fact that for nearly every state, responding to the vision

in this paper will require substantial increases in expenditures for data collection.

PURPOSES OF A STATE ASSESSMENT SYSTEM

An enhanced state assessment system should serve at least three broad purposes. When taken together, these purposes provide a compelling rationale for strengthening current activities, and provide the basis for identifying the major components of such a system as well

Policy Development

States need good information to support the development and revision of education policy. In order to develop new policy initiatives, data are required which both assesses the effectiveness of the education system in meeting state goals, and describes the nature of current education practices, processes and programs which affect goal attainment. Only when data on both outcomes and practices are available in combination, can state policy initiatives be targeted on those aspects of the system in greatest need of strengthening, and be designed to "fix" that part of the system most in need of repair.

For example, performance data which reports on the poor writing skills of secondary students can signal a problematic area in need of attention. However, in order to craft an effective policy response, data which accurately locates the source of the problem in the education system are also needed. Is the problem characteristic of all secondary students, or are particular subgroups more heavily affected than others? Is the problem largely curricular in nature, because students lack the time and opportunity to either practice writing or read the types of literature which cultivate writing skill? To what extent is the problem largely instructional, because many teachers lack the training or skill necessary to teach or evaluate student writing? Or, is the problem largely structural, in that students in some tracks are unable to imagine realistic and appropriate audiences for their writing exercises? To what extent do the mix of problems contributing to student writing performance manifest themselves differently at different grade levels, or in different

types of schools and school districts? Accurate answers to these and other questions are preconditions to the design of appropriate policies.

The need for data on educational practices and outcomes does not end with the formulation and enactment of new policy initiatives. In fact, as implementation begins, the need for sound data persists. For example, in recent years, states have enacted a bewildering array of reform initiatives, aimed at teachers and other professionals, students, school buildings, and school districts. States with comprehensive reform packages are likely to have enacted new school district accreditation standards, remedial programs, teacher evaluation requirements, career ladder or mentor teacher programs, teacher salary increases, and, perhaps a school incentive program, all with the objective of increasing student achievement. Policymakers in such states need to know how well their student achievement goal is being achieved. More specifically, they need to know how well the various programs are being implemented at the local level, what factors enhance or impede local implementation, and what type of implementation assistance is most needed. Further, they need to know what the cumulative and interactive affects of the various programs are. For example, do career ladders and school incentive programs reinforce each other's affect on collegial interactions among teachers, or do they work at cross purposes and diminish teacher professionalism? Do new state testing programs and teacher evaluation procedures combine to focus instructional activities, or do they restrict the scope of teaching and the nature of instructional practices far more than was ever intended? In addition, policymakers need to know about the cost, to the state and to local districts, of meeting new state accreditation standards? Are local funds diverted away from other, more productive purposes, such as local school improvement efforts? Do the new standards cost more or less than the remedial programs, and which delivers more with respect to student achievement gains?

The essential point here is that recent state reform initiatives are numerous, often untried, complex to implement, and likely to interact in often unpredictable ways. Without good knowledge about their cost, implementation and impact at the local level, state policymakers will have no way of knowing whether they are likely to achieve their intended effects, nor any way of

adjusting them to enhance their effectiveness and limit their unintended effects.

Beyond this, a host of fiscal factors require policymakers to have data on the implementation and impact of education reform initiatives. A host of demographic and economic factors, pressures on the federal budget, and competing priorities at the state level mean that state dollars for education will be on short supply (Kirst and Garms, 1980). Educational reform and quality cost money, and states probably don't have as much of it as they need. Consequently, in order for state policymakers to meet the policy objectives of the reform movement, they will have to spend limited state dollars as wisely as possible. Thus, the policy research questions raised above are not merely of academic interest; policymakers need to know which policies work best, and which ones cost the least, because they can't fully support all the new initiatives for very long (Kirst, 1986).

Mobilizing and Sustaining Political Support

There are a set of political needs which can be addressed by an assessment system. As Kirst (1986) has argued, the critical education policy priority for the next five years is sustaining the momentum of the education reform movement. Key to this is pursuading the public and legislators that recent investments in education are paying off. This requires evidence that recent reforms are, in fact, being implemented, and that improvements in educational quality are resulting. Consequently, states need to collect data which highlights the short- and long-term impact of education reform.

Evidence of the impact of reform is particularly critical to sustaining reform momentum for several reasons. First, despite the substantial reform and tax increase packages enacted in many states, the political coalitions supporting education reform are rather fragile. Indeed, legislative support for sustaining education reform appears to be somewhat fragile. In a number of states, legislators who lead the reform package through the legislature, and state boards responsible for implementing new legislation, are both beginning to report concern over "legislative backsliding." That is, they are concerned

that legislatures in upcoming sessions, will weaken reforms passed in recent sessions.

Further, unlike previous reform efforts (e.g., compensatory, special, vocational, and bilingual education) which attracted constituencies by adding easily identifiable programs, the current reform movement does not appear to have a natural and committed constituency it can count on for the long run. In as much of the impetus for reform originated outside the education community, educators themselves often appear to be as much a target for reform as a constituency of it. While the public has strongly supported the reform effort, it is not sufficiently organized to exercise the political clout needed. Finally, the business community has generally supported the reform movement. However, its continued support is likely to be rather dependent upon "results." Hence, the need to collect data on the implementation and impact of reform.

Improving Education Quality

Thus far we have treated assessment largely as a tool for generating information and knowledge to serve policymakers. However, state assessment systems can, and do, serve as an instrument of educational improvement well beyond informing the state policymaking process. The assessment system can be a tool available to states to more directly influence educational practice at the local level. We see at least four ways in which this purpose can be accomplished.

First, simply by virtue of what is selected to be measured and reported, a state assessment system can focus the attention of educators on important goals, outcomes and processes. For example, state testing systems have important effects on curricular and instructional decisions at the district, school building, and classroom level (Porter, 1983; Darling-Hammond and Wise, 1985). Presumably, the regular measurement and reporting of aspects of the schooling process beyond student academic performance, such as dropout and retention rates, or rates of parental and community involvement, would similarly shape local practice.

Second, by setting performance standards and attaching sanctions -- for students, teachers, or entire school buildings -- for those achievements which are measured, an assessment program can have even more powerful effects at the local level. One example of this is minimum competency testing for students. Yet another is New York State's Resource Allocation Program (RAP), in which low performing schools are required to participate in a state technical assistance program.

Third, recent years have witnessed the initiation of school incentive programs, in which individual school performance on a series of indicators is linked to the provision of financial incentives, such as discretionary dollars for individual building school improvement programs, or bonuses for staff. Such "merit school" plans, intended to foster goal consensus and collaboration on a schoolwide basis, are being tried in states such as South Carolina and Florida.

Finally, the state assessment system needs to be conceived in light of the states' regulatory functions. Historically, state accreditation of schools has been focused largely on assessing adequacy of inputs to the educational process. More recently, states have begun to focus their accreditation standards to assure the adequacy of educational practices as well. For example, states such as Mississippi, Indiana, Vermont and West Virginia have begun to develop and implement "performance-based" accreditation systems, in which school or district accreditation is linked to the use of "effective schools" practices.

This shifting focus of state regulatory systems is important for the design of state assessment systems for two reasons. First, to the extent that state assessment systems incorporate measures of educational practice as well as outcomes, there is likely to be substantial overlap between the content of the assessment system and the content of the regulatory system. At a minimum, this will require careful coordination to avoid redundancy, and may well influence the ways in which particular practices are operationalized.

Further, there is reason to suspect that school accreditation policies will continue to evolve, becoming substantively even closer to assessment

systems. There appears to be an emerging consensus, evidenced by the recent Carnegie and NGA reports, that states have been over-regulating schools and consequently as well as harmfully reducing local discretion and professionalism.

What is called for instead, these reports argue, is a more dramatic shift in the regulatory strategy employed by states. Rather than regulating inputs or processes, states should instead inspect the outcomes of schooling, accrediting only those districts or schools where outcome performance is satisfactory. Decisions regarding the level, use, and structuring of resources, as well as the specific programs and practices to be implemented, should be made at the building or district level, and remain largely unregulated by the state. To the extent that this vision is incorporated into state policy, state assessment systems will become part and parcel of the school accreditation process and will take on even greater importance. Decisions about the content and operationalization of outcome measures will have enormous influence on educational practice.

Thus far, we've argued that assessment systems can, and should, be important tools of state policy leadership in education. Well designed assessment systems can make important contributions to the policymaking process, and can directly impact on the quality and improvement of the education system in a state.

COMPONENTS OF A STATE EDUCATION ASSESSMENT SYSTEM

Guiding Principles

A state assessment system needs to serve each of the purposes discussed above. In designing such a system, several principles should be kept in mind:

1. The primary focus of the assessment system should be the school site level. The vast literature on effective schools and school improvement underscores the importance of the individual school building as a fundamental unit of improvement. For a state assessment system, this implies that the primary level at which data is reported should be the school building, and that

the reports should be designed to be useful to decisionmakers in individual buildings. Further, it implies that a major focus of assessing the implementation and impact of state reform packages should be on their cumulative and interactive effects on educational practices in buildings, more so than on the workings of discreet program components.

2. State assessment systems need to incorporate many more indicators than student academic performance. In its whole, the system should reflect:

- o the long range educational goals of the state;
- o the logic of the state's reform initiative, attending to both long- and short-term effects, as well as both intended and unintended consequences;
- o the knowledge base regarding effective schools and school improvement; and
- o public perceptions of important schooling practices and outcomes.

Only when each of these perspectives are reflected in the set of indicators employed by the state will the assessment system adequately serve its school improvement, policy development and political missions.

3. State assessment systems will need to rely on multiple measurement and data collection strategies. Some components of the system will require a complete enumeration of schools, while others will be adequately served by sampling strategies. Further, in order to obtain needed data, some mix of school records, tests, surveys, and observation and interview studies will be required. In most cases, the state should bear the full additional costs of data collection. While this will substantially increase the state's investment in assessment, the returns in terms of improved practice and policymaking will more than justify the cost. Further, the cost is likely to be only a very small portion of the state's overall investment in education.

4. State assessment systems will need to be adaptive, continuously evolving over time. The content, structure and operational procedures of a state assessment system should be modified in response to such forces as early

experience with the system, changing state goals for education, growth in the knowledge base regarding effective practices, and natural shifts in the attention of state policymakers so that as issues cycle through the policy process, the system anticipates information needs.

5. The major focus of state assessment systems should be on changes over time. The most important question an assessment system needs to answer is whether, and to what extent, the system is improving over time. In addition and especially at the school building level, it is also important to determine the relative effectiveness of each school, when compared to other similar schools.

Key Components

The major building blocks of useful state assessment systems will be Annual School Profiles for each school in the state and a set of Quality Indicators derived from samples of schools, staff, students, and the public.

Annual School Profiles

The primary purpose of the Annual School Profile is to stimulate and support school improvement efforts through the provision of useful and useable data. The profiles should provide basic, simple descriptive data on school performance and on those factors which both contribute to performance and can be altered by school staff. In this way, the profile can become the basis of the development and monitoring of individual school improvement plans. Data collection and reporting requirements need to be kept to a minimum, though the more useful the data, the higher the threshold at which a given level of respondent effort becomes a respondent burden. Further, in order for the profiles to be useful as intended, the state will need to accompany profile reporting with technical assistance and training.

The profile should indicate three broad categories of data. The first should include a set of outcome measures and might include indicators of:

- student academic performance;

- o teacher and student attendance rates;
- o drop-out and completion rates;
- o performance of students at next levels of schooling;
- o parent and student satisfaction;
- o other state goals; and
- o individual school goals, using indicators developed by each school.

Second, the profile should include measures of educational practice known to be linked to school effectiveness. These might include indicators of:

- o consensus on school goals;
- o instructional leadership, as measured by teacher perceptions;
- o opportunity to learn, as measured by such indicators as allocated time, time spent on homework, and course enrollments;
- o school climate, as measured by teacher and student perceptions;
- o teacher engagement in staff development activities; and
- o collegial interaction among teachers.

Third, the profile needs to include a set of input measures, which might include:

- o school enrollment;
- o socioeconomic, racial and ethnic composition;
- o proportion of limited English speaking, handicapped;
- o enrollments in categorical programs;
- o staff characteristics, such as education level and certification status;
- o financial resources; and
- o school expenditure patterns.

In addition, where appropriate, the school profiles could include information from accreditation programs.

The intention here is to provide each school with a fairly comprehensive and rich picture of its current performance, strengths, and weaknesses. The

indicators were selected to reflect the fact that schools pursue multiple goals, as well as the fact that goal attainment is dependent upon an interrelated set of technical practices and social conditions.

It should be acknowledged that the set of indicators goes beyond either existing data, or what is contemplated in most state or district school profile reports. The cost and effort involved in collecting the additional data is significant, or at least likely to be perceived as such by school staff. On the other hand, most of the additional data can be obtained through teacher, student and parent surveys. Such surveys are frequently part of the "needs assessment" phase of increasingly widespread effective schools programs. Thus both field tested instrumentation and experience in efficient data collection techniques typically exists for most of the proposed indicators. And, the anecdotal evidence from many effective schools projects suggests that the data collection costs are generally outweighed by their benefits.

In order for the profiles to be most useful in guiding school improvement efforts, they should be reported in a fashion which allows for a number of comparisons to be made.

First, subgroup differences in outcomes should be reported. Differences among racial, ethnic, and socioeconomic groups within the school should be reported, as well as differences among students at different achievement levels. This will focus attention on students most in need of help, and will encourage schools to reduce performance disparities as well as raise overall performance levels.

Second, data should be reported in a fashion which allows for comparisons among schools with similar inputs. Essentially, schools should be compared with other schools serving similar clientele. This reduces the likelihood that schools serving advantaged students will invariably come out on top, and increases the likelihood of success for schools with disadvantaged populations.

While this approach "evens up" the competition among schools, it will also have the effect of setting lower expectations for schools with substantial minority and low income populations. Consequently, the profiles should be

reported in a fashion which provides for comparison between the current and prior years' profile for each school, as well as for growth rates among comparable schools. These steps should serve to continuously ratchet up expectations, especially for low achieving schools.

In addition, for those measures for which it is possible, comparisons between a school's profile and state standards and goals should be made. This should stimulate improvement for all schools. When necessary, the state can readjust its standards upward.

Finally, though the annual school profiles are intended primarily for use by individual school buildings, they are also useful for local district and state policymakers as well. Both outcomes and practice measures should help local districts identify those schools most in need of help. They should also help to focus local technical assistance, staff and curriculum development, and other resources.

State policymakers can also use school profiles to target their own assistance services, as is currently the case in New York State's Resource Allocation Program. Additionally, the profiles could be incorporated into school accreditation programs and school incentive programs.

Quality Indicators

On their own, the Annual School Profiles are not adequate for all of the purposes a state education assessment system must serve. While they are well suited for stimulating local school improvement efforts, they are too limited to fill the information needs required for policy development and for mobilizing political support. What is required in addition is a broad set of Quality Indicators which reflect the overall condition, performance and direction of the state education system.

As suggested above, these indicators need to reflect four basic perspectives:

- o the long-range educational goals for the state, presumably set by the state board of education;
- o intended and unintended, short- and long-term consequences of recent state reform initiatives;
- o the growing knowledge base regarding school improvement and effective schooling practices and the overall health of the education system; and
- o public perceptions regarding important schooling characteristics and outcomes.

These four perspectives are not mutually exclusive, but each of them represents an important and useful starting point to develop indicators.

We will first explore the implications of each of these perspectives for the content of the Quality Indicators and subsequently examine their implications for data collection strategies.

Long-Term Goals

Most state boards throughout the country establish long-term goals for education; there is considerable variation among the states in how recently goals have been established, how frequently they are reviewed and updated, and how centrally they figure in the state policymaking process. Nonetheless, the state's goals are important, both because they provide direction for the education system and because they provide the basis for articulating the role of the education system in a broader social context, thereby providing a rationale for public support for education. In virtually every state, student academic achievement figures prominantly in the long-term goals. However, the goals may also address other aspects of the education system, such as enhancing student interest in independent learning, recruiting and retaining talented individuals in the teaching profession, reducing teenage pregnancy rates and drop-out rates, establishing stronger ties with the state's business community, and increasing public support for education. Indicators which reflect these goals might include, for example:

- o student mastery of subject matter and cognitive skills, as indicated by performance on standardized tests;
- o college enrollment rates within two years after high school graduation;
- o successful academic performance of college freshmen from the state;
- o success of high school graduates in entering and completing training in the military;
- o utilization of public libraries and museums by school-age youth;
- o SAT and GRE scores, and college GPAs of individuals entering the teaching profession;
- o percent of experienced teachers planning on, and actually leaving, the profession;
- o reasons for voluntary departure from the profession;
- o teenage pregnancy rates;
- o high school drop-out rates;
- o number of school-business partnership programs;
- o satisfaction of employers with recent high school graduates;
- o number of volunteers in the schools;
- o indicators of parent involvement in schools; and
- o public attitudes towards schools.

This list is simply illustrative, and would undoubtedly vary from state to state. It does serve to illustrate several key points, however. First, the set of indicators which reflect goal attainment clearly go beyond student achievement and, when considered together, they begin to provide a rich composite of how well the education system meets the goals and expectations held for it. Second, these indicators vary considerably in terms of their ease of measurement as well as measurement strategies and data sources appropriate to them. Third, they undoubtedly vary in importance as well as in the extent to which they reflect the unique contributions of the education system.

Consequences of Recent Reform Initiatives

A system of Quality Indicators needs to be sensitive to the short- and long-term, intended as well as unintended consequences of recent reform

initiatives. The need for short-term indicators results from the fact that many complex reforms, such as career ladders, will take a long time to be implemented and be reflected in student achievement gains. However, the logic of most reform initiatives does suggest a set of interim indicators which may begin to show movement long before student achievement is influenced, if they are working as expected. For example, career ladders are intended to enhance instructional effectiveness by attracting and retaining more talented individuals into the profession, by providing beginning teachers with appropriate support and mentoring, by increasing the frequency of professional and collegial interaction among teachers, and by increasing teacher participation in important professional activities, such as curriculum development, staff development, and other school improvement activities.

Even less complex reform initiatives may be slow to be reflected in student achievement gains, and short-term indicators which show their early impact can help sustain public support. For example, in most states recent increases in high school graduation requirements did not take effect until several years after their enactment, and an additional four years will transpire before the cohort they initially apply to actually graduates and reaps their cumulative achievement benefits. Thus, there will be a long time between enactment and results. In the short run, however, states can anticipate and look for changes in patterns of course enrollments and course offerings in secondary schools.

As important as short-term indicators are, indicators of potential unintended effects are equally important. For example, opponents of career ladder plans have often argued that such plans will increase competition and reduce collaboration among teachers, depress teacher morale, substantially increase teacher time on paperwork, and fail to establish credible and legitimate teacher evaluation systems. If the critics are right, this needs to be determined in the early stages of implementation, so that policies and procedures can either be adjusted or abandoned. If they are wrong, data which demonstrates this might serve to reduce opposition and build needed support as implementation proceeds. Indicators of the extent of teacher collaboration, the level of teacher morale, teacher time utilization and perceptions of the teacher evaluation system would accomplish these purposes.

Similarly, initiatives such as increased graduation requirements may have unintended effects. One obvious concern is the effect on drop-out rates of marginal students. Yet other effects are possible as well. For example, in states where the increased costs for needed additional teachers are not fully covered by the state, increased requirements may direct local funds away from other important efforts, such as remedial programs. Thus indicators of the fiscal effects of state policies on local districts would be important.

Effective Schooling Practices

This perspective will encourage states to focus on the cumulative effects of their reform packages on the overall quality of local practice. Unlike the previous starting point which has a "top down" perspective in that it starts with state-initiated reforms, an effective schools perspective looks at that point in the system where "the rubber meets the road," and asks whether the whole package of reforms is working. This is important because students, teachers and schools don't experience discreet programs in a vacuum. Rather, they experience at once the whole reform package along with a variety of local initiatives and conditions almost as a seamless web, with cumulative, interactive and generally missing and confusing results.

From this perspective, indicators of overriding importance include:

- amount of time students spend engaged in academically challenging work;
- teacher subject matter expertise;
- the extent to which teachers employ state-of-the-art instructional practices;
- the quality of teacher evaluation practices;
- the extent to which schools have climates supportive of success;
- the nature and quality of instructional leadership;
- the quality and appropriateness of inservice training available to teachers; and
- teachers' sense of professionalism, commitment and supportive work environment.

Conceptually, many of these indicators overlap with those discussed under the educational practices portion of the Annual School Profile. Conceivably a good deal of the data requirements in this area can be met with those data. However, much of the data called for here goes beyond what is feasible in the school profiles, because they require more costly and indepth data collection procedures. For example, it is difficult to measure the nature and quality of instructional practices or student engagement rates, without indepth classroom observations occurring over time. This can only be accomplished with a representative sample of teachers.

Public Perceptions of Schooling Characteristics and Outcomes

Because one of the purposes of the state assessment system is to mobilize public support for education, it is important to ensure that the Quality Indicators reflect those aspects of the schooling process that the public at large, and particular segments of the public (e.g., the business community, adults without school-age children) uses to judge the quality of public education. In all likelihood, some form of public input, such as public hearings, or surveys, may be necessary to reflect this perspective. However, one can imagine that the public would value such indicators such as:

- o student achievement;
- o teacher commitment;
- o amount of time students spend on homework;
- o frequency of discipline problems; and
- o student work attitudes, and the like.

Implications for Data Collection and Reporting

This perspective to generating Quality Indicators has several implications. First, it should be clear that the four perspectives discussed above are not mutually exclusive. The indicators suggested by each of them will overlap with those reflective of the others. This is useful, because it is one way to establish priorities among potential indicators, and reduce what could otherwise be a long list of measures. In addition, however, it will be important for states to establish priorities among potential indicators,

because it will neither be possible nor desirable to collect and report on what could be a laundry list of measures. Priorities will be based on judgements regarding the usefulness, feasibility, technical adequacy and cost of each indicator.

Second, even when priorities are established, it is neither feasible nor desirable to gather data for each indicator from all schools throughout the state. Rather, a carefully selected series of samples of schools, teachers, students, and the public would be a much more desirable approach. The samples should be selected so as to be representative of the state as a whole, as well as major subgroups one would want to generalize to (e.g., urban/suburban/rural schools, racial subgroups, experienced and novice teachers, and the like).

Third, consideration needs to be given to the tradeoffs between selecting a small number of samples from which a large amount of data is collected, or selecting a larger number of samples from each of which small amounts of data are generated. The former approach increases the analytic value of the data, while the later spreads the respondent burden more evenly. For example, if data on teacher morale and attitudes are collected from the same teachers whose classroom instruction is observed, the ability to relate the two is enhanced. On the other hand, this sample of teachers would bear an unusually heavy respondent burden.

Further, consideration needs to be given to the frequency of data collection and reporting for each indicator. While it makes sense for the state to report Quality Indicators on an annual basis, it may not be necessary for the entire set to be reported that frequently. For example, a three year cycle of surveying teachers, students and the public, one group each year, may suffice. Similarly, measures of the quality of teaching, based on costly observations, may be required only once every two to three years.

Finally, consideration should be given to alternative structures and strategies for collecting data. For example, state education department staff could be responsible for collecting some data, while other data collection efforts could be contracted out. Classroom observation data could be collected

through an agreement with one or more universities, thereby providing valuable training opportunities and possibly reducing costs as well.

With respect to the reporting of Quality Indicators several issues are important. First, as is the case with Annual School Profiles, subgroup differences in outcomes and opportunities should be reported, as well as statewide averages.

Second, the most important comparisons to be made are those which track changes and progress over time. Especially with respect to indicators of goal attainment and quality of practices, the most important concerns focus on the extent of improvement over time.

Implications for State-By-State Comparisons

This paper has been developed from a state level perspective. The assessment system proposed here is intended to serve state-level users and state needs. Those responsible for designing national assessment systems, of necessity, take a different perspective. And, for both national and state decisionmakers, the value of state-by-state comparisons on indicators of educational quality is one which warrants careful consideration at the time assessment systems are being designed.

From the perspective taken in the this paper, there is no overriding or inherent value to state-by-state comparisons on any particular set of indicators. In fact, the emphasis given to indicators which reflect state goals and state reform initiatives mitigate against across-state comparisons, at least insofar as the goals and reform initiatives differ among the states in important respects. Further, we've argued that the purposes of a state assessment system can best be served largely through comparisons of performance and practice over time, either for the state system as a whole, or for subsystems within the state. This is because an emphasis on improvement over time maximizes the chance of success for everybody, while at the same time recognizes no ceiling on performance.

Additionally, there are real limits to the inferences to be drawn from state-by-state comparisons. The likelihood of finding an appropriate comparison group for any given state, with comparable resource capabilities, student populations, political cultures and policy environments seems remote. Thus, the fairness of comparisons among states is always likely to be called into question. And, the number of states is too small and the differences among them too large to make comparative analyses, aimed at teasing out the relative effectiveness of one policy approach or another, while "controlling" for other factors, unfeasible. Comparisons among states on performance measures are unlikely to resolve whether school incentive programs are more or less effective than career ladders, or, for that matter, whether Tennessee's statewide approach to career ladders is superior to Utah's local option approach.

These observations notwithstanding, we do see enough merit to state-by-state comparisons to recommend that procedures which will make them possible be built into state and national designs. This recommendation is derived from two grounds. First, there are enough indicators likely to be common across states, to make such comparisons possible. Effective schooling practices are important to track, uniformly so among the states. Neither their value nor their operationalization needs to be dependent upon state goals or reform initiatives. In fact, they are important regardless of the particular goals or policies of a state. The same is true for such outcome measures as drop-out rates, parental and public satisfaction, or teacher morale.

The second basis for this recommendation is largely political. In skillful hands at least, state-by-state comparisons can be used to build and sustain support for education, or needed improvements in the education system. For those states that rank, or improve their ranking to, near the top, such evidence can be used to help the public understand that their schools are doing a good job, and should be recognized and supported on account of it. For those states ranking near the bottom, such evidence can used to marshall public support for needed improvements in the system. Evidence of this sort was used in states such as Mississippi, South Carolina and Arkansas to mobilize support for reform initiatives and for substantial increases in state expenditures for education. Finally, any state that finds itself ranking near the bottom on

such indicators as teacher morale, use of effective practices, student performance -- even if the rankings are not entirely fair -- ought to examine its educational system and the policies that govern it, quite carefully.

REFERENCES

Cohen, Michael. 1985. Meeting the Information Needs of State Education Policymakers: Report of a Survey of State Policymakers. Alexandria, VA: State Education Policy Consortium.

Darling-Hammond, L. and A. Wise. 1985. "Beyond Standardization: State Standards and School Improvement." The Elementary School Journal (85)3 p. 315-336.

Doyle, D.P. and T.W. Hartle. 1985. Excellence in Education: The States Take Charge. Washington, DC: American Enterprise Institute.

Fuhrman, S. and L. McDonnell. 1985. Mapping Instate Information Networks in Education Policy: An Exploratory Study. Alexandria, VA: State Education Policy Consortium.

Kirst, Michael W. 1986. "Sustaining the Momentum of State Education Reform: The Link Between Assessment and Financial Support." Phi Delta Kappan 67(2): 341-345.

Kirst, M.W. and W.I. Garms. 1980. "The Political Environment of School Finance Policy in the 1980s." In James W. Guthrie (Ed.) School Finance Policies and Practices. Cambridge, MA: Ballinger.

Porter, Andrew C. 1983. "The Role of Testing in Effective Schools." American Education. January-February 1983: 25-28.